

Valuation Property with an Automated Valuation Model /The Example of Residential Real Estate in Ulaanbaatar/

A. Enkhsuren¹, B. Enkhtuya², T. Bayartogs³, D. Zolzaya⁴, J. Munkhbileg⁵

¹Graduate School of Business, Mongolian University of Science and Technology, ULaanbaatar, Mongolia, +976

²School of Business administration and humanities, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia, +976

³School of Applied Sciences, Mongolian University of Science and Technology, ULaanbaatar, Mongolia, +976

⁴School of Information and Telecommunication Technology, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia, +976

⁵University of the Humanities, Ulaanbaatar, Mongolia, +976

Abstract— In Mongolia, real estate is assessed using traditional approaches and methods. When valuing real estate, the property's characteristics, location, and socio-economic factors are studied and the value is calculated. In this activity, the historical data of the factors influencing the real estate price and information such as international standards and legal acts are important for the accurate determination of the value of the asset. Completing a hypothetical estimate of real estate market value is time-consuming for an appraiser. With this research work, we believe that it is possible to facilitate human labor in the evaluation process, save time, avoid possible subjective factors in the calculation, and optimize the calculation based on the characteristics and location of the real estate to be evaluated and comparable assets.

Keywords— Artificial intelligence, machine learning, big data, automation, information.

I. INTRODUCTION

The digital transition and the era of artificial intelligence are beginning all over the world. At this time, we aim to save time and resources by using artificial intelligence in all fields of business, technology and science. The real estate appraisal industry has also kept up with this transition and is using the advancements in information technology for the development of the industry.

Automated Valuation Models (AVMs) are computer algorithms that use mathematical modeling and data analytics to estimate real estate values. The original idea for this automatic valuation model dates back to 1922 when George Casper Haas, a graduate student at the University of Minnesota, proposed in his master's thesis that the first automatic valuation models (AVMs) use regression to more accurately value farmland for property tax and foreclosure purposes. From this initial idea, attempts were made to use mathematical modeling in the field of property valuation, and in the 1970s, the use of regression to fully estimate the value of real estate increased.

The first trials of automated valuation models (AVMs) were conducted in the United States and Denmark. In the 1990s, the first patent for an automatic calculation model was obtained. In the late 20th and early 21st centuries, automated

valuation models (AVMs) gained popularity among appraisers.

According to the research work "Automated valuation models (AVMs): implications for the profession and their clients" conducted by the International Council of Valuation Standards (RICS) in April 2022 Automated Valuation Models (AVMs) are being used by real estate market participants around the world to evaluate residential real estate for lending, tax, and sales purposes.

The International Association of Assessing Officers, a world leader in property valuation, valuation management, real estate tax policy standards, professional development and research, has announced that "Automated Valuation Model (AVM) is based on general and special collected data from location and market. is a mathematically based computer software that calculates the market value based on market analysis of the condition and characteristics of the property. A differentiating feature of automatic valuation models is the calculation of market value generated by mathematical modeling. The reliability of the automatic valuation model (AVM) depends on the data used and the skill of the designer who created the automatic valuation model (AVM) based on the real estate data of the country or region.

In Mongolia's asset valuation sector, there is a need to study and introduce real estate valuation activities based on data using automatic valuation models, econometrics and artificial intelligence (AI) modeling techniques using the opportunities of digital transition.

II. FACTORS AFFECTING THE PRICE OF RESIDENTIAL REAL ESTATE

Factors affecting real estate prices have been studied in the automated valuation model (avms) standard for market price formation from the supply and demand side. In the automated valuation models (avms) standard (2018), iass specify the determination of real estate values based on economic, asset location, and property-related information.

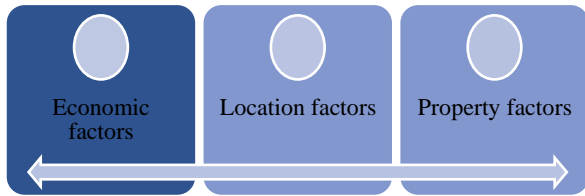


Fig. 1. Factors affecting real estate prices.

Economic factors

In the study "Economic Risk Factors and Commercial Real Estate Returns -1997" by David Lind and Andy Naranjo, macroeconomic factors such as real estate and bond market returns, interest rate volatility, and inflation rate. When determining the correlation between government events, it was mentioned that the real estate market, like the stock and bond markets, is directly related to the general economic conditions of the country.

Because there are many models that describe the factors influencing housing prices, Chen and Patel's (1998) research examines housing prices from a micro-based and macro-based approach to the housing market. In their study, the determinants of housing prices are determined from the neoclassical economic theory model, which represents macroeconomic equilibrium relationships. In the approach of the new classical economic theory, the change in housing prices is considered as a consequence of the imbalance between its supply and demand. The study considered that the demand function for housing depends on demographic factors, income, mortgage interest rates and the cost of housing, while the supply function depends on the cost of land, construction costs and the ability of construction companies to obtain loans. In 2013, Bank of Mongolia economist D. Enkhzaya studied the factors affecting housing prices in her study "Factors affecting housing price growth". In her research, she considered cost-related factors such as construction materials, workers' wages, financing costs, and land prices from the supply side, while from the demand side, she considered the ability to purchase housing with household income and housing loans.

Location factors

Real estate must be tangible, immovable, and inseparable from land. If separated, it loses its economic purpose. This quality is defined as "inseparable from the earth" by its immovable physical characteristics. So, it can be concluded that the real estate is fixed on the ground, immovable, long-term use and the environment in which it exists has a significant contribution to increase the value of the asset.

Dr. S. Dorjsuren in his book "Property Valuation Methodology" says, ".....Location is a major factor affecting real estate value. The state of a piece of land is determined by how well it meets the purpose of use, how close it is to the economic environment, population density, infrastructure development, and transportation distance. Changes in the environment in the industry can increase or decrease the value of assets. A change in the environment may cause the owner to improve his capital and incur certain expenses, or it may be a new source of income. In addition,

distance from the city center is an important indicator of location. Margorzata Rymarzak and Ewa Siemińska (Małgorzata Rymarzak and Ewa Siemińska 2012) say, "The choice of real estate location is based on a combination of various factors and criteria. When companies choose real estate to operate, they determine their location depending on the nature of the business and the specifics of the activity, which is related to the conditions considered important by the investor. However, in the case of real estate for other purposes, location decisions are determined differently from business constraints. "Determinants of location are not fixed, but change over time in response to changes in local laws and regulations and market conditions." Most studies have considered location factors to be directly related to property. Aalto University, T. N. Dim, in his research work "Automated valuation models (avms) in commercial real estate market 2019" location factors are social well-being environment, zoning, land conditions, geographic location, access to transportation, transportation convenience, parking, environmental impact, impact on government services, distance to city, district and business center.

Property factors

Property factors are the factors that most determine the value of a property based on its physical characteristics. The physical characteristics of a property vary depending on its intended use.

TABLE I. Factors affecting real estate prices.

Factors affecting real estate value		
Economic factors	Location factor	Property factor
<ul style="list-style-type: none"> • Employment rate • Unemployment • Inflation • Interest rate • Gross domestic product • Exchange rates • Mortgage interest • Demographic growth • Household income and expenditure • Rising prices in the construction industry • Increase in land prices • Other /Other considerations related to the type of real estate/ 	<ul style="list-style-type: none"> • Population density and concentration • Distance from urban areas • Distance to government and business establishments • Distance to health facilities • Distance to educational institution • Access to public transport • Road communication • Distance to airport and railway station • Infrastructure • Other / Other considerations related to property type 	<ul style="list-style-type: none"> • Zoning • Construction design • Infrastructure • Type of building use /residential, office, industrial, etc./ • Date of construction • Construction period • Total area /assessed area/ • Total floors • Accommodation floors • Number of rooms • Number of toilets • Warehouse • Window size and location • Elevator • Building planning and elevation • Quality of building use • Car parking • Quality of land use • Ground position • Land size • Other /Other considerations related to the type of real estate/

M. Mooya (Manya M. Mooya 2016) mentions in his real estate valuation theory that "The price changes not only

depending on the type of real estate, but also on its physical characteristics." He mentioned the number of rooms, building design, land size, floors, etc. as examples of physical characteristics.

Barry and Rodriguez (Barry and Rodriguez 2004) studied the impact of building characteristics on the return on investment in residential construction in India. As a result of the study, it was determined that the design, materials, interior design, layout, number of rooms, and area utilization of residential buildings have a significant impact on the return on investment and increase the value.

III. MODEL THEORETICAL APPROACHES TO AUTOMATIC VALUATION

Most automated valuation models (avms) estimate the value of an asset over a period of time by analyzing the values of comparable properties. The purpose of automatic appraisal is to determine the market value of public and private property in a short period of time.

Automated valuation models (avms) cannot accurately model large asset impairments, abnormal conditions of market distress, excessive leverage, and excessive operational lag. Therefore, when choosing this model, appraisers can choose which asset to use, and the appraiser can make adjustments by taking into account the large-scale damage, uniqueness, and difficulty of the asset. It has advantages over the mass evaluation method.

The theoretical methodology of the automatic valuation model is traditional for asset valuation¹:

- market approach,
- income approach,
- cost approach.

In the automatic assessment model standard, the assessment methodology is developed based on the above 3 main assessment approaches.

Cost approach: cost approach methods are indirect methods of estimating market value. Replacement cost is calculated based on market value less depreciation. Adjusts the cost approach method by reviewing construction costs and sales. This method provides more acceptable results for new properties, special properties, and properties that are not selling well. Land and buildings are determined separately. The following formulas are used to value real estate using automated valuation models and the cost approach.

$$MV = \omega GQ \times [(1 - BQ_D) \times RCN + LV]$$

MV- calculation of market value;

ωGQ - general quality variables /location, time, etc./;

BQ_D - quality variables /depreciation calculation/;

RCN- rehabilitation/replacement cost;

LV- land value

$$MV = \omega GQ * [(\omega BQ \times \sum BA) + (\omega LQ \times \sum LA) + \sum OA]$$

MV- calculation of market value;

ωGQ - general quality variables;

ωBQ - product of building quality variables;

$\sum BA$ - sum of additional building variables;

ωLQ - product of land quality variables;

$\sum LA$ - sum of additional land variables;

$\sum OA$ - sum of additional variables.

Based on the cost information provided by the customer, the value is determined by making adjustments based on the market data during the evaluation using an automatic valuation model.

Income Approach: The variables involved in the income model are income per unit (net rental area), cost per unit or capitalization rate. The steps in this method are:

1. Calculate the net amount of total income and expenses,
2. Choose the appropriate capitalization method (model definition),
3. Multiplier for calculating capitalization rate or income (model calibration),
4. Determining value by capitalization.

$$MV = NOI/R$$

MV- calculation of market value;

NOI- net cash flow from operations;

R- capitalization rate.

Another approach to the income approach uses the relationship between gross income and rent. These are called gross income multiplier (GIM) and gross rent multiplier (GRM).

$$MV = GI_A * GIM \quad \text{or} \quad MV = GI_M * GRM$$

MV- calculation of market value;

GI_A - annual gross income;

GI_M - gross monthly income;

GIM- gross income multiplier;

GRM- Gross Rent Multiplier.

Market approach, Sales Comparison Method: In the sales comparison method, the valuation process is performed in two steps. The merge step measures the disparity using a series of data filters to select the most comparable sales for the object.

The second step compares the sale price of comparable properties for sale to the object based on the differences in data attributes. The adjusted sales price will be used to determine the value of the property.

$$MV_s = SP_c + ADJ_c$$

MV_s - calculation of market value;

SP_c - represents the comparable sales price / property for sale/;

ADJ_c - displays comparable coefficients /sales/.

Direct market approach: The direct market method chooses one of the three forms of this model to predict market value. These are linear additive model, multiplicative model, mixed-non-linear hybrid model.

Additive model:

$$MV = B_0 + B_1 \times X_1 + B_2 \times X_2 \dots \dots \dots$$

MV- calculation of market value or relevant variable;

B_0 - formula;

B_j - related percentages or coefficients;

X_j - independent variable.

In the direct market method, "MV" refers to the unit selling price or total selling price.

Multiplicative model: The results are amplified by increasing the coefficients found in the model.

$$MV = B_0 \times X_1^{B_1} \times X_2^{B_2} \times \dots \dots \dots$$

- MV- calculation of market value or relevant variable;
- B₀- formula;
- X_i^{B_i}- X independent variable;
- X_i^{B_i}- B_i related rates or coefficients.

A multivariate multiplicative model consists of a constant B₀ and an associated adjustment. Multiplicative models are usually fitted using a set of linear regressions.

Hybrid model: Mixed (non-linear) models are a combination of additive and multiplicative models. Value is determined by building, land and "other" components (eg outbuildings).

$$MV = \omega GQ \times \left[(\omega BQ \times \sum BA) + (\omega LQ \times \sum LA) + \sum OA \right]$$

- MV- calculation of market value;
- ωGQ- general quality variables;
- ωBQ- build quality scrapers;
- ∑BA- total additional variables of construction;
- ωLQ- land quality variables;
- ∑LA- total additional land variables;
- ∑OA- total additional variables.

Automated valuation models are all based on market comparison methods, but are most effective when this large

amount of sales data is available. the market comparison method is based on the principle of substitution. The essence of the substitution principle lies in finding the difference between the best reference market price and the profitability of the asset being valued.

IV. MACHINE LEARNING METHODOLOGY

Machine learning is a subset of artificial intelligence and is primarily concerned with algorithm development. Arthur samuel defined machine learning as "...the ability for machines to automatically learn from data, improve experimental results, and predict things without specific programming...".

A machine learning method has a model defined by several parameters. Machine learning is the execution of a program to optimize the parameters of a model using data or accumulated experience. A model can be predictive to determine future predictions or descriptive in order to gain knowledge from data. Machine learning techniques use statistical theory to build mathematical models.

Machine learning is classified as follows depending on the techniques and data used.

TABLE II. Machine learning methods.

No	Teaching method	Main content
1	Supervised learning	The basic idea is that your data consists of examples of situations and examples that describe results. The machine then uses the training data to build a model that can predict new outcomes based on past examples. If the outcome categorical column is qualitative data, the model is "classification" and if it is quantitative data, the model is "regression".
2	Unsupervised learning	Because the training data does not have an undefined "outcome", we cannot easily evaluate the output of this algorithm in the same way as with supervised learning. Therefore, the training will be found by classifying the data into similar and different qualities and creating clusters.
3	Semi-supervised learning	It can be said that it is an intermediate course between Supervised and Unsupervised training. The training has mostly labeled data and some data is unknown. The goal of the unsupervised model is to classify some unlabeled data using a labeled dataset.
4	Active learning	The model chooses a label (category) to combine which operations and data benefits. In this way, if you can create an algorithm to learn from the results of the data you want to learn using the design methodology of very small classified data, then the training can be better than the traditional methods.
5	Reinforcement learning	Build a state model and learn from it every time you make the right choice. Based on accumulated experience, you will learn what to do in a given situation in order to achieve the desired goal over a certain period of time. Back-to-back learning is related to control theory, Markov decision processes, and game theory.

In real estate valuation, decision trees and random forest methods are widely used in real estate valuation using databases, automatic valuation methods, and machine learning to perform valuation work. Because real estate prices are constantly changing due to the influence of the market.

Decision tree

A decision tree algorithm starts by choosing the function that best splits the data. The "best partition" is often defined by criteria based on metrics such as the Gini mixture or data growth in regression tasks. At each node of the decision tree, the selected feature is used to partition the data into subgroups. For example, the selected feature is location. Starting from the root node, a decision node and a recursive split are created to select the most influential feature. Each leaf node represents a prediction of the value of a particular property based on a combination of features. This interpretive model helps to understand the reasoning behind the assessment. Evaluation measures such as absolute mean error confirm the prediction.

Entropy is used to identify distinct patterns in data. If the data is uniform, the entropy is 0, and if the entropy is 1, the data is divided into two equal parts. For example, when calculating the entropy of data with one attribute or target entropy, use the data frequency table as follows2.

$$E(S) = \sum_{i=1}^c -p_i \log_2 p_i$$

If "pi" is the frequency or probability of element/category "i" in our data. If there are two types of attributes, entropy is calculated using the frequency table.

$$E(T, X) = \sum_{c \in X} P(c)E(c)$$

Information augmentation relies on entropy reduction after partitioning data on attributes. Building a decision tree involves finding features that maximize information gain. Information augmentation is calculated by calculating the total entropy as a proportion of each branch of the tree minus the partition entropy.

$$\text{Gain}_{T,X} = \text{Entropy}_T - \text{Entropy}(T,X)$$

The first node of the decision tree, or the root node, is selected as the decision node with the attribute with the highest information gain. The Outlook feature has the greatest data growth. After selecting the feature with the highest information gain as a decision node, divide the data into branches and repeat the same process for each branch. If the entropy is 0, then the tree is a leaf. Branches with entropy greater than 0 should be split further.

Random forest

The "Random Forest" algorithm is to create a combination of decision trees. Each tree predicts the value of a property based on different characteristics. Through a process called bagging (bootstrap aggregation), multiple decision trees are trained on different subsets of the data. An algorithm combines these predictions to create a reliable and accurate

overall estimate. Random forests reduce overfitting by introducing randomness into feature selection and dataset sampling, improving model generalizability.

$$g_{RF}(x) = \frac{1}{B} \sum_{b=1}^B g_{T_b^*}(x)$$

B- the number of decision tree

$g_{T_b^*}(x)$ -trained decision tree

V. VALUATION OF RESIDENTIAL PROPERTY USING AN AUTOMATED VALUATION MODEL (AVMS)

Automated Valuation Models (AVMs) use automated valuation models (AVMs) to value properties using direct market approaches and sales comparison methods. The market approach and sales comparison method of asset valuation is done by comparing the valuation item to the comparable asset.

TABLE III. Machine learning methods.

No	Indicators	Classification	Evaluation criteria
1	Zoning	1- AA	5 km radius from the city center
		2- A	8 km south-east and south-west from the city center
		3- B	6-10 km from the city center
		4- C	More than 11 km from the city center
2	Year	1- Very old	Before 1970
		2- Old	1971-2000
		3- Kind of new	2001-2018
		4- New	After 2019
3	Area size	1- Small	Less than 40m2
		2- Medium sized	41-80 m2
		3- Big sized	More than 81m2
4	The number of rooms	1- Small	Between one to 2 rooms
		2- Medium	Between 3 to 4 rooms
		3- Big	More than 4 rooms
5	The location of the floors	1 On top	10- 10th floor and up
		2 In the middle	Between 5th and 9th floors
		3 At the base	Between 1st and 4th floors
6	Discounts enjoyed after negotiation	1- much	20 to 30 Percent discount
		2 - medium	10 to 19 percent discount
		3- small	Zero to 9 percent discount
7	Construction design	1- A	Reinforced concrete cast constructions, cast concrete paving /completely cast/
		2- B	Reinforced concrete structure with brick and block filler and poured concrete pavement
		3- C	Precast reinforced concrete
		4- D	Mixed (concrete belt and core) construction, load-bearing brick, block filler, precast, cast pavement
		5- E	Load bearing and wood composite construction, mixed aggregates and wood, wood and lightweight board pavement
		6- S	Two-sided metallized panels /sandwiches with foam insulation and light construction, wood and lightened panels, two-sided metallized panels /sandwiches with foam insulation, wood and lightened panels, two-sided metallized panels /sandwiches with foam insulation
8	Engineering network	1- A	Unconnected, with normal stove heating
		2- B	Semi-connected, Low-pressure cooker
		3- C	Fully connected
9	Distance to kindergartens, schools, service centers and transport networks	1- Far	More than 5 kms
		2- Medium	2-3 kms
		3- Near	1-2 kms

Random Forest Algorithm for Automatic Real Estate Appraisal is a combination of decision trees. Each tree predicts a property's value based on each characteristic of the factors that influence price. After that, the random forest algorithm combines the joint predictions of the decision tree to create a reliable and accurate overall estimate.

Using the original data of the property data, each node of the random forest is estimated by the explanatory variables and the results are combined to make a final conclusion.

This process ensures that every decision is made and the trees grown may be different and incomparable. The final prediction of the target variable is a weighted vote, or the sum of the predicted averages. The present value of the property is determined by this sum.

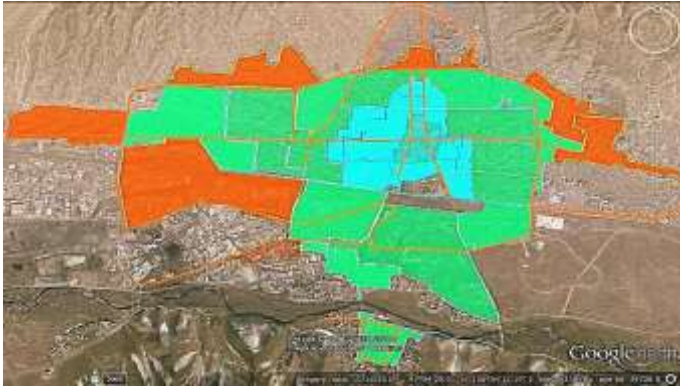


Fig. 2. Real estate zone map of Ulaanbaatar city

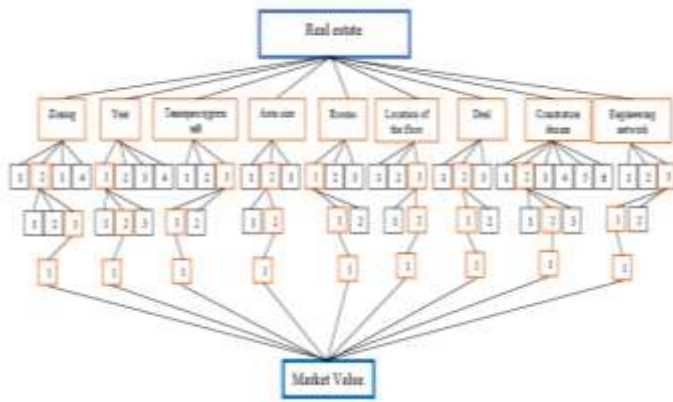


Fig. 3. Valuation of residential property using machine learning using decision trees

VI. CONCLUSION

It was observed during the research that the research and development of automatic evaluation methods has become more and more active in recent years. For example, the developed countries of the world are widely using the automatic valuation model when valuing real estate for mortgage, tax, trade, and insurance purposes, creating time and cost savings. On the other hand, appraisers specializing in the evaluation of intellectual property, business, and extractive property focus more on putting assets into economic circulation and attracting investment. Internationally, the next generation of automated valuation models has been defined by using artificial intelligence to evaluate real estate based on real

estate photos, evaluating data points using the resulting data, and combining quantitative and qualitative indicators.

Developing an automated assessment model to assess some assessment items can save time and money.

For example: real estate taxes, bank loan collateral, litigation, and insurance appraisals.

It is concluded that it is possible to create an asset valuation database in Mongolia and evaluate assets using automatic valuation models and machine learning methods.

REFERENCES

- [1]. Standard on Automated Valuation Models, 2018
- [2]. Standard on Mass Appraisal of Real Property, 2017
- [3]. Published by Australian Property Institute, "Australia and New Zealand Valuation and Property Standards" 2012
- [4]. Bogdan Trawinski, Wrocław Piwowarczyk, Zbigniew Telec, Wrocław University of Science and Technology "Comparison of expert algorithms with machine learning models for real estate appraisal", 2017
- [5]. A. Geenen H. Nguyen R.T. Wiersma, "Automating Valuations for Real-Estate", 2017
- [6]. Tuğba GÜNEŞ and Ümit YILDIZ, Turkey, "Mass Valuation Techniques Used in Land Registry and Cadastre Modernization Project of Republic of TURKEY", 2015
- [7]. Nils Kok, Eija-Leena Koponen, and Carmen Adriana, Martínez-barbosa "Big Data in Real Estate? From Manual Appraisal to Automated Valuation" 2017
- [8]. Professor Mike Hefferan, University of the Sunshine Coast and Adjunct Professor Terry Boyd, Central Queensland University and University of the Sunshine Coast, "Property Taxation and Mass Appraisal Valuations in Australia – Adapting to a New Environment" 2010
- [9]. Brano GLUMAC, François Des Rosiers, "Real estate and land property automated valuation systems: A taxonomy and conceptual model" 2018
- [10]. Michael Bekech, Chair, "Computer Assisted Mass Appraisal Systems" 2002
- [11]. Tengxiang Su, "A BIM and Machine Learning Integration Framework for Automated Property Valuation", 2022
- [12]. Hilgers, B.A.J, "Automated Valuation Models for Commercial Real Estate in the Netherlands", 2018
- [13]. International valuation standards., 2020
- [14]. Manya M. Mooya., Real Estate Valuation Theory., 2016
- [15]. Appraisal Institute., The Appraisal of Real Estate 15th edition., 2020
- [16]. <https://www.rics.org/globalassets/rics-website/media/knowledge/research/insights/future-of-valuations-insights-paper-rics.pdf>
- [17]. RICS., Automated valuation models (AVMs): Implications for the profession and their clients., 2022
- [18]. S. Dorjsuren., Valuation of Immovable properties., 2005
- [19]. Profession evaluation institution., Basics of Property Valuation., 2014
- [20]. G. Gantulga., Real Estate Relations 2011
- [21]. D.Zolzaya, B. Tssetsgee., Basic of Artificial Intelligence., 2021